

Vocal Impact of a Prolonged Reading Task at Two Intensity Levels: Objective Measurements and Subjective Self-Ratings

*Angélique Remacle, †,‡Camille Finck, *Anne Roche, and *Dominique Morsomme, *†Liège, and ‡Mons, Belgium

Summary: Objectives. The objectives of this study are to evaluate the impact on voice of both duration and intensity level of 2 hours of continuous oral reading. Voice modifications accompanying changes in intensity level during prolonged reading tasks are analyzed.

Methods. Fifty normophonic women undergo two sessions of voice loading in which the required intensity level of voice varied between 60–65 dB(A) for the first session and 70–75 dB(A) for the second session. The effects of loading on objective data (average fundamental frequency [F0], jitter%, shimmer%, noise-to-harmonic ratio, maximum phonation time, lowest frequency [F-Low], highest frequency [F-High], frequency range [Range], lowest intensity [I-Low] level, and highest intensity level) and self-ratings (voice quality, phonation effort, vocal fatigue, and laryngeal discomfort) are assessed every 30 minutes during the loading tasks.

Results. Results indicate that average F0, F-Low, I-Low, maximum phonation time, feeling of phonation effort, vocal fatigue, and laryngeal discomfort increase during prolonged reading, whereas shimmer% and self-rating of voice quality decrease. Average F0, F-High, and Range are the only parameters influenced by the required intensity of vocal load; they are significantly higher in the 70- to 75-dB session compared with the 60- to 65-dB session. Concerning the subjective self-ratings, similar results for the four ratings used suggest that only one would suffice in future studies.

Conclusions. These results confirm the importance of both duration and intensity level as loading factors, even if intensity level affects fewer variables than duration.

Key Words: Vocal loading—Loading factors—Voice level—Acoustic analysis—Subjective ratings.

INTRODUCTION

Vocal load is the acoustic vocal power integrated over time,¹ depending on the amount of voicing. The duration and intensity level of the voice are two known loading factors. Effects of vocal load have been examined in several experiments, differing from one to another in terms of loading task and experimental environment. Some studies have observed the effect of vocal load using a natural speech material,^{2–9} whereas others have used a loading task as prolonged loud reading^{10–24} or repeated vowel sequence.²⁵ The duration and intensity level of such loading tasks have varied across studies as well as the moment and duration of recordings. Voice samples have been analyzed either throughout the vocal load using audiotape recorders^{4,7,14} and dosimeters or at different times before and after loading.^{2,3,5–7,10–13,15–17,19–25} Some experiments have taken place in a laboratory environment,^{10–24} whereas others have studied subject's vocal load, during the working day.^{2–9}

The present study has been initiated to improve the understanding of voice duration and intensity level as loading factors. It is part of a larger project that addresses occupational vocal load in teachers. Indeed, teaching is a vocally demanding profession that requires a prolonged use of voice at

a high intensity (HI) level, in an elevated background noise.⁴ The main loading factors in teaching voice are the duration and intensity level of voice required, which are risk factors for developing an occupational voice disorder.²⁶ Therefore, voice problems occur more frequently in teachers compared with the general population.^{27–30}

Effects of duration as voice loading factor

The duration of vocalization corresponds to the *time dose* or *voicing time*.^{14,31} Many studies have investigated the effect of prolonged vocal use, focusing on the duration as a main loading factor. Generally, such studies have analyzed the voice before and after vocal load, either in a laboratory or on field. Given the various methodologies and different intensity levels required for the vocal task, cross-study comparisons are difficult. Videostroboscopic examinations have not shown systematic consistent after vocal load changes in glottis configuration.^{10,12,13,21} Recent studies using high-speed digital imaging with phonovideos have identified vibration behavior changes in healthy voices concerning closing and opening dynamics and more left-right vocal fold asymmetries.^{2,15} Acoustically, the majority of the studies have demonstrated a raise in fundamental frequency (F0)^{3–7,9,10,19,21} and intensity^{5,6,19} of voice and a decrease in instability values (jitter and shimmer).^{6,21} Some authors^{5,7} have considered these results as a normal adaptation to loading. Subjectively, previous studies have shown a significant increase in self-rated voice symptoms during the working day^{3,6,9} or after a loading task performed in a laboratory.^{10,16,17,32}

Effects of intensity level as voice loading factor

Voice intensity changes with the amplitude of vocal fold's oscillations, influencing the *distance dose* and the *dissipated*

Accepted for publication July 29, 2011.

From the *Department of Psychology: Cognition and Behavior, University of Liège, Liège, Belgium; †ENT Department, Centre Hospitalier Universitaire de Liège, Liège, Belgium; and the ‡Faculté de Psychologie et des Sciences de l'Éducation, University of Mons, Mons, Belgium.

Address correspondence and reprint requests to Angélique Remacle, Department of Psychology: Cognition and Behavior, Unité de Logopédie de la Voix, University of Liège, Rue de l'Aunaie, 30 (B38), 4000 Liège, Belgium. E-mail: Angelique.Remacle@ulg.ac.be

Journal of Voice, Vol. ■, No. ■, pp. 1–10

0892-1997/\$36.00

© 2011 The Voice Foundation

doi:10.1016/j.jvoice.2011.07.016

energy dose.^{14,31} Voice intensity affects the shear stress in the ligament³³ and the medial compression of the vocal folds (compressive stress).³⁴ Mechanical stress and thus injury risk is positively correlated with pressure, proportional to the voice intensity.³⁴ Factors impacting voice intensity are background noise, room acoustics, number of listeners, and distance between the speaker and listeners.²⁹ Voice intensity generally increases in noisy conditions,^{29,35} known as the Lombard effect. Speaking in background noise constitutes a risk factor for vocal load³⁶ and for developing occupational voice disorders.^{26,29,37} Moreover, an increased vocal intensity is usually associated with an F0 elevation.³⁸ Nevertheless, by investigating preschool teachers in their normal workplace, Linstrom et al⁸ found no general correlations between noise exposure and both vocal intensity level and average F0. These findings suggest that vocal behavior in relation to noise exposure is individual.

To investigate the effects of intensity level, previous studies^{35,36,39} observed the voice production during realistic noise exposure using a method that enables cancellation of the background noise from the speech signal in five different noise conditions. One result is that having to make oneself heard over the noise resulted in increased vocal intensity, F0, and phonation time. Subjectively, women reported less success in making themselves heard and higher effort than men.³⁵ Both patients with vocal nodules and controls increased their self-ratings of strain as an effect of the increased background noise level.³⁶

Effects of both duration and intensity level on voice

Few studies have observed the effect of both intensity level and duration of the loading task. Vintturi et al^{18,32} studied the changes in voice during a vocal loading test with, among others, a specific emphasis on the speech output level. Neils and Yairi⁴⁰ examined the influence of various levels and durations of noise exposure in six females reading in noise during 45 minutes.

The current work aims to evaluate the impact of both duration and intensity level of vocal load on voice in normophonic women, with no exposure to background noise. The following questions are examined: (1) how does the voice vary with the duration of vocal load and (2) how does the voice vary with two different intensity levels of vocal load? First, we hypothesize that a prolonged vocal use would modify objective measures and subjective self-ratings of voice. Second, by comparing low and high intensity levels, we expected that objective measures and subjective self-ratings would be different. Another goal of this study is to examine if there is any difference in the results obtained in the four subjective self-rating scales.

METHODS

Subjects

Fifty females (mean age = 25.4 years, SD = 4.98, range = 21–47) were recruited as subjects. All of them met the following criteria: no laryngeal pathologies established by anamnesis and videolaryngostroboscopic examination, no complaint or history of voice disorders, no voice therapy, no hearing disorders, no upper respiratory infection at the time of the study, no profession or hobby that implied frequent voice use, and no voice

training. All subjects were nonsmokers. Eleven of the subjects had some vocal education in their childhood (drama or singing lessons). According to the experimenters' perceptual judgment, subjects had normal voice on the days of the experiment.

On the two days before each testing, subjects were asked to avoid any vocally abusive behavior (singing, loud talking, shouting, and yelling). They received the instructions to sleep and drink normally and not to ingest caffeine, alcohol, or any medication that causes drying of the vocal folds. All subjects provided informed consent but were blind to the study hypothesis. They received an oral and a written explanation of the experimental procedure.

Procedure

Loading task. Subjects' voices were orally loaded by reading a novel of their choice for 2 hours. Each subject underwent two sessions in which the intensity level of reading varied. The first session required a low intensity (LI) level, between 60 and 65 dB(A). The second session required a high intensity (HI) level, between 70 and 75 dB(A). A mean interval of 22 days was kept between the two sessions, with a minimum of 5 days allowing the recovery from the first reading session.

During the loading task, participants were seated in a quiet room (background level < 30 dB(A)) and instructed to read aloud. Voice intensity level was constantly controlled with a digital sound level meter (Velleman, DVM805, China) at a distance of 40 cm from the mouth. The examiner encouraged the participants to maintain the intensity level if it differed from the target level. The relative humidity of ambient air was controlled using a hygrometer (DOSTMANN electronic, P600, Wertheim-Reicholzheim, Germany) and kept constant (30% ± 10%). The reading was filmed. Intensity control and recordings were made by the first and third authors.

Evaluation protocol. For each reading session, serial sets of objective measurements and subject self-ratings were carried out every 30 minutes: (time 0 = T0) before the loading task, (time 1 = T1) after 30 minutes of reading, (time 2 = T2) after 1 hour of reading, (time 3 = T3) after 1 hour and 30 minutes of reading, and (time 4 = T4) after 2 hours of reading. The protocols for the two reading sessions are reported in Table 1. All measurements were repeated using an identical protocol.

Every 30 minutes, during the reading task, the researcher advised the participants to drink one glass of water, to ensure that they remain hydrated. For each reading session, the entire procedure (loading tasks and evaluations) took 3 hours per subject.

Objective measurements. Serial voice quality objective data were obtained by the use of *Multi-Dimensional Voice* program (MDVP) (Kay Elemetrics, Lincoln Park, NJ): average F0, jitter% (Jitt), shimmer% (Shim), and noise harmonic ratio (NHR). Three samples of the (sustained) vowel /a/ at a comfortable pitch and intensity level were analyzed and then averaged for a final value. Each production was recorded with a constant mouth-to-microphone distance of 7 cm.

The lowest frequency (F-Low), the highest frequency (F-High), the frequency range (Range), the lowest intensity

TABLE 1.
Protocols for the Two Reading Sessions

LI Session (60–65 dB)	HI Session (70–75 dB)
Objective measurements and self-rating (LI0)	Objective measurements and self-rating (HI0)
Reading session (30 min)	Reading session (30 min)
Objective measurements and self-rating (LI1)	Objective measurements and self-rating (HI1)
Reading session (30 min)	Reading session (30 min)
Objective measurements and self-rating (LI2)	Objective measurements and self-rating (HI2)
Reading session (30 min)	Reading session (30 min)
Objective measurements and self-rating (LI3)	Objective measurements and self-rating (HI3)
Reading session (30 min)	Reading session (30 min)
Objective measurements and self-rating (LI4)	Objective measurements and self-rating (HI4)

(I-Low), and the highest intensity (I-High) were collected with the *voice range profile* program (VRP) (Kay Elemetrics, Lincoln Park, NJ) on the vowel /a/. F-Low, F-High, and Range were recorded during three trials, the subject gliding from a middle range note to the lowest possible note and then to the highest possible note. I-Low and I-High were collected at c1 pitch (262 Hz), which is in the middle of an estimated female frequency range. Subjects were asked to sustain the target pitch at the softest and the loudest possible level three times successively. During the recording session, subjects could observe their performance on the monitor. To motivate participants to perform at their maximum capacity, the investigator provided verbal encouragements and auditory examples if necessary.

Objective measurements involved also the determination of maximum phonation time (MPT), which is a measure of the vocal function in relation to the glottis efficiency and voice quality. Each subject was asked to produce three samples of the prolonged vowel /a/ at comfortable pitch and intensity, as long as possible, in normal voice. The longest phonation time obtained was analyzed.

Subjective self-ratings. To know how the subjects cope with the vocal loading, they were asked to answer the following questions using a 100-mm horizontal visual analog scale (VAS) every 30 minutes: (1) how is your voice quality (how does the voice sound)? The extremes on the VAS were 0% for a poor voice quality and 100% for a normal voice quality; (2) do you feel any phonation effort (strain or effort to produce the voice)? The extremes on the VAS were 0% for no vocal effort and 100% for a maximum vocal effort; (3) do you feel any vocal fatigue (tiredness of voice or in neck muscles)? The extremes on the VAS were 0% for no vocal fatigue and 100% for a maximum vocal fatigue; and (4) do you feel any laryngeal discomfort (pain or dryness in your throat)? The extremes on the VAS were 0% for no laryngeal discomfort and 100% for a maximum laryngeal discomfort. VAS is a commonly used measurement instrument for self-rating of subjective characteristics related to voice.^{3,6,9,17,32,35,36} The continuous aspect of this scale provides more subtle differences than a discrete scale. The state-

ments used in the present study were based on the questionnaires presented in the literature.^{3,6,9,10,16,17,32,35,36} Before the loading task, subjects were informed that they will have to complete the same questions every 30 minutes. When subjects scored their complaints, they did not have access to their previous ratings.

Statistics

A two-way repeated-measures analysis of variance (ANOVA five duration \times two intensity level) where the subjects ($N = 50$) were used as their own controls was carried out to compare data obtained at different times of the reading sessions. When the ANOVA showed significant differences ($P < 0.05$), a post hoc honestly significant difference (HSD) Tukey was computed either on the main significant effects or on the interaction, to compare the means. When necessary, for meeting the assumption of homogeneity of variances and normality, logarithmic transformations normalized raw data before ANOVA. For clarity, means of the raw data values are presented in figures. All calculations were conducted using the statistical software *Statistica/Win* (version 8.1, StatSoft Inc, Tulsa, OK). Figure 1 shows the results derived from the objective measurements. Figure 2 shows the results obtained from the subjective self-ratings.

An ANOVA was used to analyze if there were any statistically significant differences in subjective self-ratings at each time, for each reading session. When the ANOVA showed significant differences ($P < 0.05$), a post hoc HSD Tukey was computed.

RESULTS

Objective measurements

For F0 (Figure 1A), results from ANOVA revealed a significant main effect of the duration ($F(4,196) = 12.5$, $P < 0.0001$). Post hoc comparisons showed an increase in F0 between T0 and T1, T2, T3, T4, as well as between T1 and T4. At each time, F0 measurements for HI session were higher than for LI session, as a main effect of the intensity level ($F(1,49) = 21.9$, $P < 0.0001$). No interaction between duration and intensity ($F(4,196) = 0.3$, $P = 0.90$) was demonstrated.

Results for Jitt (Figure 1B) showed no significant main effect of the duration ($F(4,196) = 1.64$, $P = 0.17$), no significant effect of the intensity level ($F(1,49) = 0.69$, $P = 0.41$), and no interaction between duration and intensity ($F(4,196) = 0.69$, $P = 0.60$).

Regarding Shim (Figure 1C), a significant main effect of the duration was found ($F(4,196) = 6.62$, $P < 0.0001$). The post hoc test showed significant differences between T0 and T1, T2, T3, T4. There was no significant main effect of the intensity level ($F(1, 49) = 0.12$, $P = 0.73$). No interaction between duration and intensity ($F(4,196) = 0.31$, $P = 0.87$) was found.

Concerning the NHR (Figure 1D), there was no significant main effect of the duration ($F(4,196) = 1.82$, $P = 0.12$) and no significant main effect of the intensity level ($F(1,49) = 0.10$, $P = 0.75$). No interaction between duration and intensity ($F(4,196) = 0.05$, $P = 0.99$) was found.

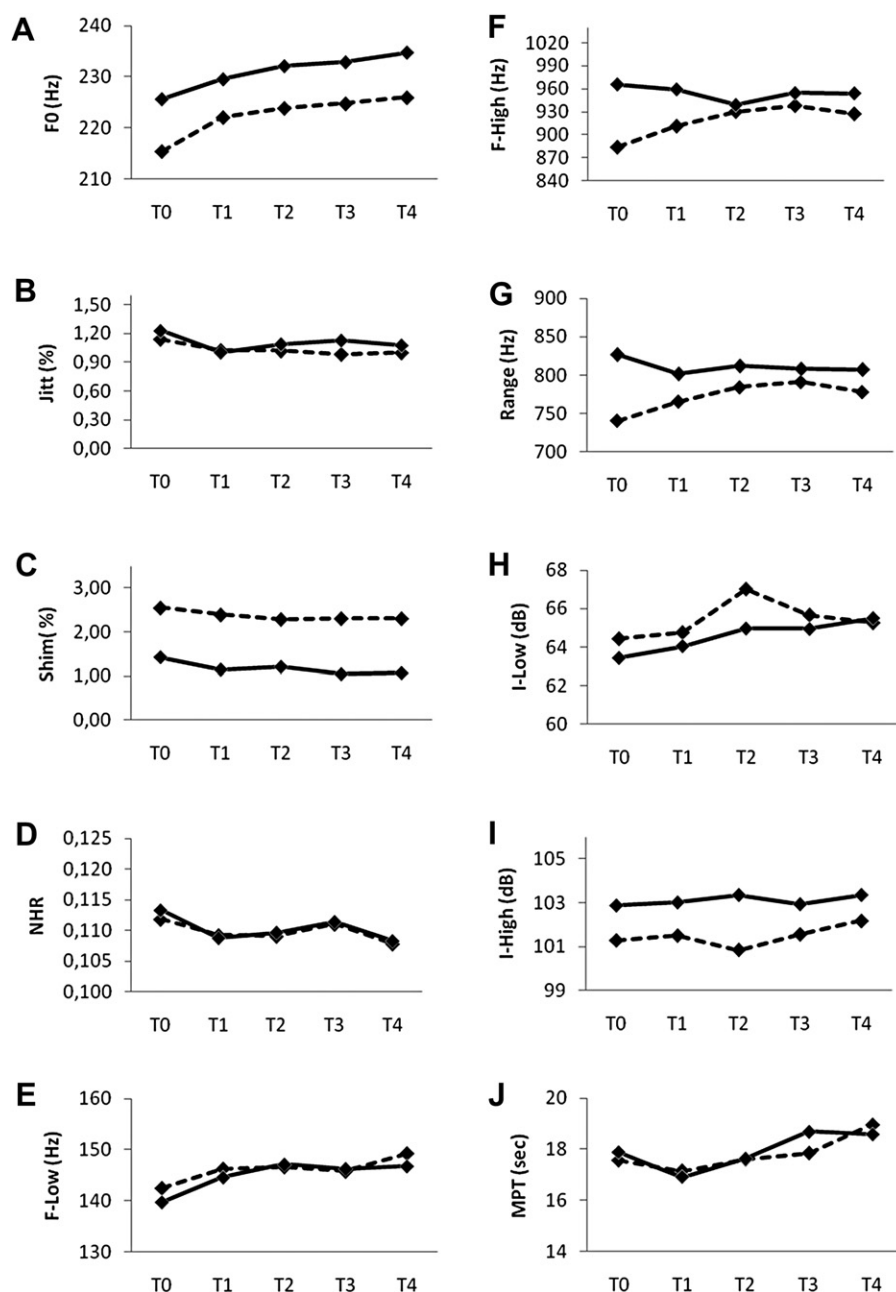


FIGURE 1. A–J. Results from the objective measurements at each time (T0, before the loading task; T1, after 30 minutes; T2, after 1 hour; T3, after 1 hour and 30 minutes; and T4, after 2 hours) for the LI session (dotted line) and HI session (solid line) during the reading task.

F-Low (Figure 1E) increased significantly as a main effect of the duration ($F(4,196) = 4.81$, $P = 0.001$). The post hoc test showed a significant increase in F-Low between T0 and T2, T3, T4. There was neither significant main effect of the intensity level ($F(1,49) = 0.31$, $P = 0.58$) nor interaction between duration and intensity ($F(4,196) = 0.46$, $P = 0.76$).

For F-High, no main effect of duration was found ($F(4,196) = 0.74$, $P = 0.56$), but there was a significant main effect of the intensity level ($F(1,49) = 7.21$, $P = 0.009$). Figure 1F shows that F-High was systematically higher for the HI session than for the LI session. There was no interaction between duration and intensity ($F(4,196) = 2.08$, $P = 0.085$).

For the frequency Range (Figure 1G), there was no significant main effect of the duration ($F(4,196) = 0.70$, $P = 0.59$) and no interaction between duration and intensity ($F(4,196) = 2.30$, $P = 0.061$). At each time, the Range for HI session was higher than for LI session, as a main effect of the intensity level ($F(1,49) = 12.94$, $P = 0.007$).

Results for I-Low (Figure 1H) demonstrated a significant main effect of the duration ($F(4,196) = 5.00$, $P = 0.007$). The post hoc test showed a significant rise between T0 and T2, T3, T4 as well as between T1 and T2. No main effect of the intensity level ($F(1,49) = 3.60$, $P = 0.062$) was observed. There was no interaction between duration and intensity ($F(4,196) = 1.3$, $P = 0.29$).

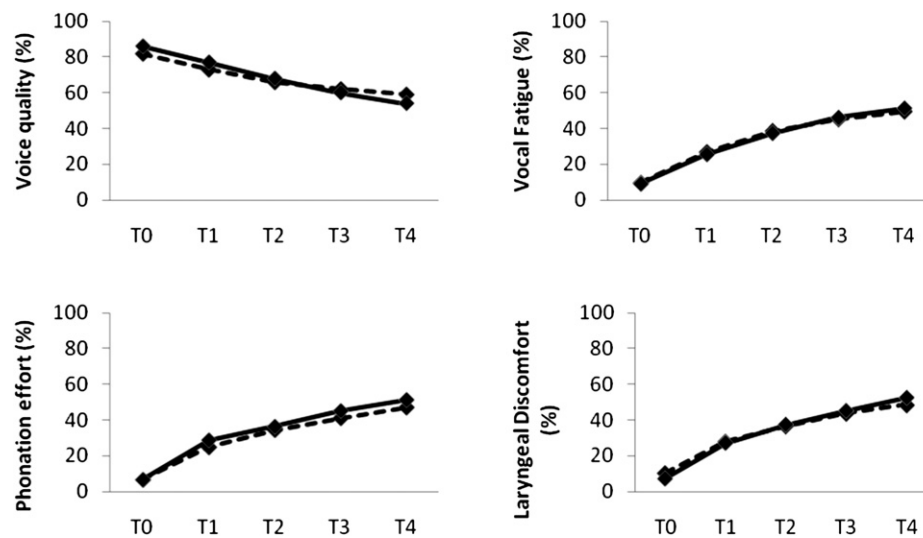


FIGURE 2. Results from the subjective self-ratings at each time (T0, before the loading task; T1, after 30 minutes; T2, after 1 hour; T3, after 1 hour and 30 minutes; and T4, after 2 hours) for the LI session (dotted line) and HI session (solid line) during the reading task.

Concerning I-High, we didn't observe any significant main effect of the duration ($F(4,196) = 0.85$, $P = 0.49$) nor main effect of the intensity level ($F(1,49) = 2.29$, $P = 0.14$). Figure 1I showed systematically slightly higher values for the HI session, but the statistical tests failed to reveal any significant difference. There was no interaction between duration and intensity ($F(4,196) = 0.78$, $P = 0.54$).

For MPT (Figure 1J), the ANOVA analysis demonstrated a significant main effect of the duration ($F(4,196) = 4.37$, $P < 0.01$). The post hoc test showed a significant rise between T1 and T3 as well as between T1 and T4. No main effect of the intensity level effect ($F(4, 49) = 0.11$, $P = 0.74$) and no interaction between duration and intensity ($F(4,196) = 0.51$, $P = 0.72$) were observed.

Subjective self-ratings

Results from repeated-measures ANOVA (Table 2) demonstrated a significant effect of the duration ($P < 0.0001$) for all the subjective self-ratings, but no differences between the two

intensity levels of reading. The detailed results from the post hoc test on the main effect of the duration are presented in Table 3. A significant interaction between duration and intensity was found for voice quality (Table 4 for the post hoc). In other words, subjects reported that their voice quality worsened significantly during the reading while phonation effort, vocal fatigue, and laryngeal discomfort increased. From T0 to T2, subjects noted a better voice quality for the HI session than for the LI session, thereafter it was the opposite.

The simple ANOVA (Table 5) did not show any significant difference between the four self-rating measurements, excepted at T0 in the LI session ($F(3,196) = 4.95$, $P = 0.002$). The post hoc test demonstrated a significant difference between voice quality and phonation effort ($P = 0.0012$), as well as between voice quality and vocal fatigue ($P = 0.032$).

DISCUSSION

This study addresses how objective measurements and subjective self-ratings reflect effects of duration and intensity level as

TABLE 2.
Results From the Repeated-Measures ANOVA for Subjective Self-Ratings

Variables	Duration Effect		Intensity Level Effect		Interaction Effect (Duration × Intensity)	
	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>
Voice quality	55.66	<0.0001*	0.064	0.08	4.88	<0.001*
Phonation effort	98.56	<0.0001*	1.84	0.18	0.59	0.67
Vocal fatigue	130.55	<0.0001*	0.0005	0.98	0.47	0.75
Laryngeal discomfort	93.19	<0.0001*	0.012	0.91	1.44	0.22

Notes: Degree of freedom for the duration effect = (4,196); degree of freedom for the intensity level effect = (1,49); degree of freedom for the interaction between duration and intensity level = (4,196). *F* represents the *F* value from the ANOVA; *P* represents the *P* value from the repeated-measures ANOVA where $P < 0.05$ indicates statistical significance, shown by an asterisk.

TABLE 3.
Results From the HSD Tukey Post Hoc for Subjective Self-Ratings of Phonation Effort, Vocal Fatigue, and Laryngeal Discomfort

Variables	Time (a)	Time (b)	P	LI Session	HI Session
				Mean Difference (b – a)	Mean Difference (b – a)
Phonation effort	0	1	<0.0001*	18	22
	0	2	<0.0001*	27	30
	0	3	<0.0001*	34	38
	0	4	<0.0001*	40	45
	1	2	0.024*	9	8
	1	3	<0.0001*	16	16
	1	4	<0.0001*	22	23
	2	3	0.0001*	7	8
	2	4	<0.0001*	13	15
	3	4	0.079	6	7
Vocal fatigue	0	1	<0.0001*	17	17
	0	2	<0.0001*	29	28
	0	3	<0.0001*	35	37
	0	4	<0.0001*	39	42
	1	2	<0.0001*	12	11
	1	3	<0.0001*	18	20
	1	4	<0.0001*	22	25
	2	3	0.002*	6	9
	2	4	<0.0001*	10	14
	3	4	0.134	4	5
Laryngeal discomfort	0	1	<0.0001*	17	20
	0	2	<0.0001*	26	30
	0	3	<0.0001*	33	38
	0	4	<0.0001*	38	45
	1	2	0.0007*	9	10
	1	3	<0.0001*	16	18
	1	4	<0.0001*	21	25
	2	3	0.018*	7	8
	2	4	<0.0001*	12	15
	3	4	0.087	5	7

Note: Time 0, before the loading task; Time 1, after 30 minutes; Time 2, after 1 hour; Time 3, after 1 hour and 30 minutes; Time 4, after 2 hours. Column *P* represents *P*-values from the HSD Tukey post hoc test, where $P < 0.05$ indicates statistical significance, shown by an asterisk.

loading factors. To the best of our knowledge, this study is the first one to analyze the voice impact of a prolonged 2-hour reading task at two controlled different intensity levels in the same group of subjects. Neils and Yairi⁴⁰ used a similar design where six subjects completed a 45-minute reading task in three different background noise levels (50 dB(A), 70 dB(A), and 90 dB(A)), presented into earphones to induce different levels of vocal effort and vocal fatigue. Unfortunately, voice intensity was not controlled during the reading. Voices were analyzed perceptually, acoustically, and aerodynamically. No significant change was found, either because of the loading task duration or the background noise level. However, the results reflected a wide range of variations from subject to subject. Vinturri et al^{18,32} carried out an objective analysis on 20 females reading at LI level (<65 dB SPL) and 20 other females reading at HI level (>65 dB SPL). Each subject participated only once in the experiment. As far as the objective analysis is concerned,¹⁸ inverse-filtered data, Voice

Range Profile, and the singer's formant were analyzed during a 45-minute vocal loading session. The results showed that the loading task at HI is associated with more hyperfunctional changes than LI. As far as the subjective self-rating is concerned,³² the loading-related symptoms were observed during five sessions of 45-minute reading. Vinturri's results demonstrated that most symptoms increased significantly to a peak mean value after three or four sessions with no clear effect of the output level. Compared with other studies on duration and intensity level effects,^{18,32,40} the advantages of the current investigation are the constantly controlled voice level and the quite large group of subjects (50) implicated in both intensity level sessions. The fact that each subject was implicated in both sessions reduced the intersubject differences when studying the intensity level loading factor. As acoustic measures (Jitt, Shim, and NHR) are known to vary with intensity and frequency,⁴¹ and between subjects, it seems essential to use subjects as their own control for studying the intensity level

TABLE 4.
Results From the HSD Tukey Post Hoc Test for Subjective Self-Rating of Voice Quality

Time (a)	Time (b)	LI Session		HI Session	
		Mean Difference (b – a)	P	Mean Difference (b – a)	P
0	1	–9	<0.0001*	–9	0.0001*
0	2	–16	<0.0001*	–18	<0.0001*
0	3	–20	<0.0001*	–26	<0.0001*
0	4	–23	<0.0001*	–32	<0.0001*
1	2	–7	0.043*	–9	0.0001*
1	3	–11	<0.0001*	–17	<0.0001*
1	4	–14	<0.0001*	–23	<0.0001*
2	3	–4	0.606	–8	0.001*
2	4	–7	0.009*	–14	<0.0001*
3	4	–3	0.811	–6	0.098

Note: Time 0, before the loading task; Time 1, after 30 minutes; Time 2, after 1 hour; Time 3, after 1 hour and 30 minutes; Time 4, after 2 hours. Columns *P* represents *P* values from the HSD Tukey post hoc test, where $P < 0.05$ indicates statistical significance, shown by an asterisk.

effect. Another innovation in the present study design is the analysis of Jitt, Shim, NHR, and MPT, which were not included in previous investigations on the intensity level effects of vocal load.^{18,32,35,36,39,40}

In the literature, reading tasks last from a few seconds^{35,36,38,39} to a few hours,^{19,20,22,32} between 60¹² and 80 dB.^{10,15,21,25} Niebudek-Bogusz et al¹³ fail to demonstrate any significant videolaryngostroboscopic changes after a 30-minute reading through white noise at 80 dB SPL, as well as Gelfer et al¹² after 1 hour at 60–70 dB. Gelfer et al¹² report that “loud reading for 15 minutes or even 1 hour is not enough to induce notable laryngeal alterations in most subjects” and that “2 hours of loud reading may be the minimum required to induce laryngeal changes.” However, Niebudek-Bogusz et al¹³ do demonstrate significant acoustical changes after 30 minutes through white noise at 80 dB SPL, as well as Laukkanen et al¹⁷ after 5 minutes at 70 dB. In the current investigation, 2 hours of duration was chosen, based on different studies^{15,16,21} that demonstrated effects of vocal load using the same duration. Concerning the intensity level of the reading, previous papers^{18,19,32} report two different intensity levels of reading: <65 and >65 dB at a distance of 2 m. In the present study, 60–65 dB and 70–75 dB at a distance of 40 cm from the mouth were chosen to have a clear and controlled difference between the two sessions.

This investigation focuses on female subjects because they are known to be more sensitive to vocal load than males,^{35,42} and have a higher prevalence of voice disorders.^{27,30} The prolonged reading task took place in controlled laboratory conditions, despite of the fact that these experimental conditions do not reflect the real-life vocal use. For the HI reading session, subjects are asked to raise their vocal intensity level without any rational reason, although in real-life situations, vocal intensity generally increases as a consequence of the environmental noise. Some studies^{35,36,39} analyze the speaker’s voice during vocal exposure to background noise that represents more realistic conditions. Nevertheless, these studies did not address the voice duration. Despite the artificial setting of voice use,

the advantage of the experimental conditions used in the current investigation for studying vocal load is that factors such as loading material, duration of phonation, intensity level of voice, room acoustic, background noise, and humidity were controlled. The question of how to generalize the results to real-life situations remains a challenge. The following discussion of our observations is organized according to the two issues studied, namely: (1) how does the voice vary with the duration of vocal load and (2) how does the voice vary with two different intensity levels of vocal load? Thereafter, the similarity between the subjective self-ratings is discussed.

Effects of duration

The first hypothesis of this experiment is that voice should be modified during prolonged reading, as a duration effect of vocal load. This supposition follows from previous studies reporting objective and subjective changes in voice after a prolonged vocal use. Results for both intensity level sessions show expected modifications of objective measurements and subjective self-ratings. F0, F-Low, I-Low, MPT, feeling of phonation effort,

TABLE 5.
Results from the Simple ANOVA for the Differences Between the Subjective Self-Ratings of Voice Quality, Phonation Effort, Vocal Fatigue, and Laryngeal Discomfort

Time	LI Session		HI Session	
	F	P	F	P
0	4.95	0.002*	2.35	0.074
1	0.27	0.844	0.93	0.427
2	0.56	0.642	0.73	0.532
3	1.06	0.367	0.71	0.547
4	1.20	0.312	0.64	0.590

Notes: Degree of freedom = (3,196). *F* represents the *F* value from the ANOVA; *P* represents the *P* value from the ANOVA where $P < 0.05$ indicates statistical significance, shown by an asterisk.

vocal fatigue, and laryngeal discomfort significantly increase during prolonged oral reading, whereas Shim and self-rating of voice quality significantly decrease. Jitt, NHR, F-High, Range, and I-High do not statistically change during the reading task.

A significant rise in F0 is reported in both laboratory^{10,19,21} and field conditions^{3-7,9} as a result of the vocal load. This may indicate an adaptation to loading, as suggested by other authors.^{5,7} In their study, using electromyography to observe the effects of vocal load on laryngeal muscles, Boucher and Ayad¹¹ show that variation of F0 does not consistently reflect fatigue in laryngeal structures. In the current investigation, a significant increase in F0 can be already noted after 30 minutes, while subjects report phonation effort, vocal fatigue, and laryngeal discomfort values remaining quite low (between 25% and 29%). Thereafter, F0 continues to rise significantly over time. After 2 hours of reading, subjects estimate their phonation effort, vocal fatigue, and laryngeal discomfort between 47% and 52%. Subjects' ratings indicate a shift in the voice self-perception, but they do not note high values on the VAS. For this reason, the hypothesis that F0 rise reflects an adequate physiological adaptation of the vocal apparatus to loading^{5,7} rather than a sign of vocal fatigue is supported.³

To date, few studies have investigated vocal load effects on F-Low, F-High, and Range. Similar to the present study, Stemple et al²¹ show no significant change in the frequency range after a 2-hour reading task at 75–80 dB on a series of 10 subjects. However, the lowest and the highest pitch values tend to rise posttest, while subjects complain of the difficulty in producing the lowest pitch. In the present study, the increased F-Low after 1 hour may suggest some effect of the long duration.

When looking at Jitt, Shim, and NHR mean values, we can see that Jitt varies between 0.98% and 1.23% (MDVP standards: $0.633\% \pm 0.351\%$). The overall Jitt values are not within MDVP normal limits, perhaps because MDVP standards do not concern a French-speaking population. It does not mean that a pathology is present. Shim varies between 1.04% and 2.55% (MDVP standards: $1.997\% \pm 0.791\%$). NHR varies between 0.108 and 0.113 (MDVP standards: 0.112 ± 0.009). Shim and NHR mean values in both intensity level sessions are within normal limit values, which was expectable given that the subjects of the study are normophonic. The repeated-measures ANOVA shows that frequency instability (Jitt) and NHR do not change while intensity instability (Shim) significantly decreases as a consequence of the vocal load. Stemple et al's results²¹ demonstrate a significant decrease in Jitt after a 2-hour reading task in normophonic women. Laukkanen et al⁶ also report a significant drop in Jitt and Shim values after a working day, in female teachers. More intensity stability in Laukkanen and in the present study as well as more frequency stability in Stemple and Laukkanen studies seem to reflect adequate adaptations to loading. Conversely, Niebudek-Bogusz et al¹³ show a significantly increased Jitter in dysphonic teachers after a 30-minute loud reading. Furthermore, Shimmer increases significantly in the hyperfunctional group of dysphonic teachers after the reading task. These results show that the effects of vocal load differ from dysphonic to normophonic

subjects. Vocal stability in dysphonic subjects is more affected by the vocal load, suggesting that they are less prone to have a healthy adaptation to vocal load.

In terms of intensity measures, I-Low significantly increases in both LI and HI reading sessions, whereas I-High remains unchanged. In 40 female subjects, Vintturi et al¹⁸ report that the I-Low of the VRP rises significantly at 262 and 330 Hz after a 45-minute reading. In a study by Akerlund,²⁴ the lower phonetogram contour significantly rises in female and male dysphonic patients after reading in a 80-dB-SPL white noise for 15 minutes, although no significant movement of the upper contour is noted. Sihvo and Sala's study²² observes the phonetogram of 10 females through five sessions of 45-minutes reading during 1 day. Results show a significant increase in both fortissimo and pianissimo curves during the loading. The aforementioned studies agree that the lowest intensity rises as an effect of vocal load. The increased I-Low in the present study after 1 hour suggests some effect of the long duration. The explanation can be a modification of the viscoelastic characteristics of the vocal folds that become stiffer and drier after prolonged vocal use, resulting in an increased subglottal pressure.

In the present study, I-High does not change significantly suggesting no effect of the duration. Vilkman et al¹⁹ note that females are able to produce the highest intensity value during the last loading sample of the day. Their interpretation is that there is no true laryngeal or respiratory muscle weakness because of fatigue after a prolonged (5×45 -minute) reading task.

As far as MPT is concerned, we observe an increase from T1. The increased MPT through the time may represent an adaptation of the voice to the duration. Another explanation could be that subjects improve their performance on the task because of a training effect. To better explain the evolution of MPT, aerodynamic measurements should have been obtained. Unfortunately, it was impossible to collect aerodynamic data at the time of our study. Stemple et al²¹ and Kelchner et al¹⁰ do not show any significant modification in MPT after a 2-hour reading task at 75–80 dB, neither in normal women nor in normal adolescent males.

Finally, subjects' self-rating of voice significantly worsens throughout the reading task. These results are in agreement with previous investigations. In the present study, the voice quality already decreases after 30 minutes of reading. Kelchner et al¹⁰ observe a worsening of the voice quality after a 2-hour reading task in peripubescent boys. Lehto et al^{3,9} also report a lower voice quality (increased perceived hoarseness) after 1 working day in customer advisors. In agreement with Chang's results,¹⁶ the current work demonstrates an increased phonatory effort after a 2-hour reading task. In the present study, vocal fatigue increases through time (already after 30 minutes). Laukkanen et al¹⁷ show an increased tiredness in the throat after 45 minutes of reading at 70 dB. On field, more vocal fatigue is reported after 1 working day in customer advisors^{3,9} and more tiredness in the throat in female teachers. Our results demonstrate an increased laryngeal discomfort (including pain) through the reading task, already after 30 minutes. Laukkanen et al¹⁷ also find an increased feeling of pain in the throat after a 45-minute reading task.

Effects of intensity level

The second hypothesis of this research is that the intensity level of vocal load would have an effect on the observed parameters. Our results reveal that F0, F-High, and Range are the only parameters influenced by the required intensity of the reading task. One explanation is that speaking loudly does not necessarily imply a wrong usage of voice or a degradation of subjective sensations, especially in normophonic subjects. Another explanation is that aerodynamic measures would be more adequate to demonstrate the influence of the intensity level of voice as loading factor. This is because voice intensity is regulated by means of subglottal pressure and transglottal airflow rate that influence vibrational amplitude of the vocal folds and laryngeal tension.

F0 and F-High are significantly higher in the HI session compared with the LI session. Even if an increased F0 is a well-known strategy to cope with the vocal load, it is not without risks for vocal health because of the mechanical stress³³ applied to the vocal folds. The higher the number of vocal folds oscillatory cycles over time (*cycle dose*), the greater the number of compressive and collisional stresses that the vocal folds' cover needs to endure. Otherwise, the *distance dose* and the *energy dissipation dose* are most likely to be large with a higher intensity level. Svec et al^{14,31} suggest that the heat dissipated in the vocal folds probably does not play an important role in tissue damage, but they consider the dose of vibrations as hazardous.

Comparison of the two reading sessions demonstrates systematic higher F0 and F-High as well as an improved Range in the HI condition, even before the reading task (T0). The difference in F0 because of the intensity is well predicted from other studies looking at the relationship between F0 and intensity.^{35,38} Surprisingly, there is no consistency across the two baselines sessions. The fact that this discrepancy is great in T0 already is somewhat problematic. It is possible that (1) subjects increased the intensity and consequently the F0 of voice before starting the HI reading task that could be their own strategy for getting ready to perform the task at HI, or (2) the pattern was caused by chance. Aerodynamic measurements would again give us some insight to explain this observation. The improved performances in F-High and Range for the HI session in T0 can be the result of an effect of learning. In fact, subjects systematically underwent the LI session before the HI to preserve a progression order of vocal intensity. We wanted to be sure those subjects were able to perform the 60–65 dB reading task before doing the 70–75 dB one.

Subjectively, subjects do not report more complaints in the HI than in the LI session. One possible explanation is that subjects did not have access to their previous ratings when they scored their complaints. According to Vinturi et al,³² the output level of vocal load has no clear effect on subjective symptoms during five sessions of 45-minute reading. The reason for this might be that the difference between the two output level groups (<65 dB SPL and >65 dB SPL) in their study is not very large. In other studies,^{35,36} different aspects of voice production during realistic noise exposure were investigated. Authors used a method that enables cancellation of the background noise from the speech signal, allowing acoustic analysis

of the speaker's voice during vocal load against background noise. Using this method, 23 normal subjects completed an 80- to 110-seconds reading task in realistic environmental noise, in five different noise conditions.^{35,39} The differences between the reading conditions are closer controlled. Results show that subjects who have to make themselves heard over the noise raise their intensity level and experience more effort level and fatigue in elevated background noise conditions. In the present study, the differences between the two intensity levels are distinct and continually controlled.

Similarity of the subjective self-ratings

In the present study, we used four subjective self-ratings based on the questionnaires presented in the literature: voice quality, phonation effort, vocal fatigue, and laryngeal discomfort. An ANOVA was computed to determine how similar the results were for all the self-ratings. Results show a difference at LI0, between voice quality and phonation effort, as well as between voice quality and vocal fatigue. These differences concern only a minor part of the self-ratings, realized at LI0, which is the first time the subjects completed the VAS, before the reading task. Thereafter, the subjects do not report different results at the four VAS that may suggest that one self-rating would suffice in the future for studying the impact of vocal loading.

Clinical implications

In terms of security level of voice, it is not possible to recommend a benchmark for safe vocal use on the basis of these results. Further studies are necessary to establish safety standards concerning the duration and intensity level loading factors to avoid overloading. Nevertheless, the present study supports the recommendation not to have a prolonged voice use and a high intensity level, at the risk of developing loading related impairments such as vocal nodules or polyps. Indeed, duration and intensity level may be phonotraumatic because they both raise the F0, and thus the repeated vibrations of the vocal folds. The high-velocity impact between vocal folds during speech could induce mechanical stress related to structural disruptions of the basement membrane of the vocal folds.^{33,34} Additionally, computational models predict that mechanical stress level and thus injury risk is positively correlated with driving force that is proportional to voice intensity.³⁴

CONCLUSIONS

The present study aims to evaluate the impact of both duration and intensity level of vocal load on voice. Our results confirm the importance of both factors in vocal load, even if intensity level affects fewer variables than duration. The significant effects of duration are an increase in F0, F-Low, I-Low, MPT, feeling of phonation effort, vocal fatigue, and laryngeal discomfort, as well as a decrease in Shim and self-rating of voice quality. Increased F0, improved MPT, and decreased Shim seem to reflect an adaptation to the duration, whereas increased F-Low and I-Low after 1 hour as well as worsening of self-ratings suggest an effect because of the long duration. Concerning intensity level, F0 and F-High are significantly higher and the frequency

range larger in the HI session compared with the LI session. These results support the recommendation not to have a prolonged voice use and a high intensity level at the risk of developing loading related impairments because of mechanical stress on the vocal folds. Nevertheless, aerodynamic measurements would be required to better explain and support these results. Moreover, an analysis of the postural and respiratory adaptation to vocal load would be of great interest, as well as perceptual analysis. In future research, it would be pertinent to examine the recovery time after such vocal loading tasks; it is possible that if the intensity level is higher, the recovery time after vocal load could be longer.

Acknowledgments

The authors gratefully acknowledge the subjects of the experiments, Professor Ruth Epstein for revising the English language, Professor Etienne Quertemont for statistical advice and Dr Lionel Lejeune for assistance with the videolaryngostroboscopy.

REFERENCES

1. Titze IR. Criteria for occupational risk in vocalization. In: Dejonckere PH, ed. *Occupational Voice: Care and Cure*. The Hague, The Netherlands: Kugler; 2001:1–10.
2. Doellinger M, Lohscheller J, McWhorter A, Kunduk M. Variability of normal vocal fold dynamics for different vocal loading in one healthy subject investigated by phonovibrograms. *J Voice*. 2009;23:175–181.
3. Lehto L, Laaksonen L, Vilkmán E, Alku P. Occupational voice complaints and objective acoustic measurements—do they correlate? *Logoped Phoniatr Vocol*. 2006;31:147–152.
4. Sodersten M, Granqvist S, Hammarberg B, Szabo A. Vocal behavior and vocal loading factors for preschool teachers at work studied with binaural DAT recordings. *J Voice*. 2002;16:356–371.
5. Jonsdottir V, Laukkanen A-M, Vilkmán E. Changes in teachers' speech during a working day with and without electric sound amplification. *Folia Phoniatr Logop*. 2002;54:282–287.
6. Laukkanen A-M, Ilomaki I, Leppanen K, Vilkmán E. Acoustic measures and self-reports of vocal fatigue by female teachers. *J Voice*. 2008;22:283–289.
7. Rantala L, Vilkmán E, Bloigu R. Voice changes during work: subjective complaints and objective measurements for female primary and secondary schoolteachers. *J Voice*. 2002;16:344–355.
8. Lindstrom F, Wayne KP, Sodersten M, McAllister A, Ternström S. Observations of the relationship between noise exposure and preschool teacher voice usage in day care center environments. *J Voice*. 2011;25:166–172.
9. Lehto L, Laaksonen L, Vilkmán E, Alku P. Changes in objective acoustic measurements and subjective voice complaints in call center customer-service advisors during one working day. *J Voice*. 2008;22:164–177.
10. Kelchner LN, Toner MM, Lee L. Effects of prolonged loud reading on normal adolescent male voices. *Lang Speech Hear Serv Sch*. 2006;37:96–103.
11. Boucher VJ, Ayad T. Physiological attributes of vocal fatigue and their acoustic effects: a synthesis of findings for a criterion-based prevention of acquired voice disorders. *J Voice*. 2010;24:324–336.
12. Gelfer MP, Andrews ML, Schmidt CP. Documenting laryngeal change following prolonged loud reading. A videostroboscopic study. *J Voice*. 1996;10:368–377.
13. Niebudek-Bogusz E, Kotylo P, Sliwiska-Kowalska M. Evaluation of voice acoustic parameters related to the vocal-loading test in professionally active teachers with dysphonia. *Int J Occup Med Environ Health*. 2007;20:25–30.
14. Titze IR, Svec JG, Popolo PS. Vocal dose measures: quantifying accumulated vibration exposure in vocal fold tissues. *J Speech Lang Hear Res*. 2003;46:919–932.
15. Lohscheller J, Doellinger M, McWhorter AJ, Kunduk M. Preliminary study on the quantitative analysis of vocal loading effects on vocal fold dynamics using phonovibrograms. *Ann Otol Rhinol Laryngol*. 2008;117:484–493.
16. Chang A, Karnell MP. Perceived phonatory effort and phonation threshold pressure across a prolonged voice loading task: a study of vocal fatigue. *J Voice*. 2004;18:454–466.
17. Laukkanen A-M, Jarvinen K, Artkoski M, et al. Changes in voice and subjective sensations during a 45-min vocal loading test in female subjects with vocal training. *Folia Phoniatr Logop*. 2004;56:335–346.
18. Vinturi J, Alku P, Lauri ER, Sala E, Sihvo M, Vilkmán E. Objective analysis of vocal warm-up with special reference to ergonomic factors. *J Voice*. 2001;15:36–53.
19. Vilkmán E, Lauri ER, Alku P, Sala E, Sihvo M. Effects of prolonged oral reading on F0, SPL, subglottal pressure and amplitude characteristics of glottal flow waveforms. *J Voice*. 1999;13:303–312.
20. Sihvo M, Laippala P, Sala E. A study of repeated measures of softest and loudest phonations. *J Voice*. 2000;14:161–169.
21. Stemple JC, Stanley J, Lee L. Objective measures of voice production in normal subjects following prolonged voice use. *J Voice*. 1995;9:127–133.
22. Sihvo M, Sala E. Sound level variation findings for pianissimo and fortissimo phonations in repeated measurements. *J Voice*. 1996;10:262–268.
23. Lauri ER, Alku P, Vilkmán E, Sala E, Sihvo M. Effects of prolonged oral reading on time-based glottal flow waveform parameters with special reference to gender differences. *Folia Phoniatr Logop*. 1997;49:234–246.
24. Akerlund L. Phonetograms before and after exposure to noise. *Scand J Log Phon*. 1993;18:93–97.
25. Jilek C, Marienhagen J, Hacki T. Vocal stability in functional dysphonic versus healthy voices at different times of voice loading. *J Voice*. 2004;18:443–453.
26. Chen SH, Chiang S-C, Chung Y-M, Hsiao L-C, Hsiao T-Y. Risk factors and effects of voice problems for teachers. *J Voice*. 2010;24:183–190.
27. Villanueva-Reyes A. Voice disorders in the metropolitan area of San Juan, Puerto Rico, Spain: profiles of occupational groups. *J Voice*. 2011;25:83–87.
28. Sliwiska-Kowalska M, Niebudek-Bogusz E, Fiszor M, Los-Spychalska T, Kotylo P, Szurowska-Przygocka B, Modrzewska M. The prevalence and risk factors for occupational voice disorders in teachers. *Folia Phoniatr Logop*. 2006;58:85–101.
29. Vilkmán E. Occupational safety and health aspects of voice and speech professions. *Folia Phoniatr Logop*. 2004;56:220–253.
30. Roy N, Merrill RM, Thibeault S, Parsa RA, Gray SD, Smith EM. Prevalence of voice disorders in teachers and the general population. *J Speech Lang Hear Res*. 2004;47:281–293.
31. Svec JG, Popolo PS, Titze IR. Measurement of vocal doses in speech: experimental procedure and signal processing. *Logoped Phoniatr Vocol*. 2003;28:181–192.
32. Vinturi J, Alku P, Sala E, Sihvo M, Vilkmán E. Loading-related subjective symptoms during a vocal loading test with special reference to gender and some ergonomic factors. *Folia Phoniatr Logop*. 2003;55:55–69.
33. Titze IR. Mechanical stress in phonation. *J Voice*. 1994;8:99–105.
34. Gunter HE. Modeling mechanical stresses as a factor in the etiology of benign vocal fold lesions. *J Biomech*. 2004;37:1119–1124.
35. Sodersten M, Ternstrom S, Bohman M. Loud speech in realistic environmental noise: phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *J Voice*. 2005;19:29–46.
36. Aronsson C, Bohman M, Ternström S, Södersten M. Loud voice during environmental noise exposure in patients with vocal nodules. *Logoped Phoniatr Vocol*. 2007;32:60–70.
37. Vilkmán E. Voice problems at work: a challenge for occupational safety and health arrangement. *Folia Phoniatr Logop*. 2000;52:120–125.
38. Gramming P, Sundberg J, Ternström S, Leanderson R, Perkins W. Relationship between changes in voice pitch and loudness. *J Voice*. 1988;2:118–126.
39. Ternström S, Bohman M, Södersten M. Loud speech over noise: some spectral attributes, with gender differences. *J Acoust Soc Am*. 2006;119:1648–1665.
40. Neils LR, Yairi E. Effects of speaking in noise on vocal fatigue and vocal recovery. *Folia Phoniatr (Basel)*. 1987;39:104–112.
41. Gelfer MP. Fundamental frequency, intensity, and vowel selection: effects on measures of phonatory stability. *J Speech Hear Res*. 1995;38:1189–1198.
42. Vinturi J, Alku P, Lauri ER, Sala E, Sihvo M, Vilkmán E. The effects of post-loading rest on acoustic parameters with special reference to gender and ergonomic factors. *Folia Phoniatr Logop*. 2001;53:338–350.